PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/62, 15/83, C07K 14/08, 14/435, C12N 7/01, 5/10

A1

(11) International Publication Number: WO 96/12027

(43) International Publication Date: 25 April 1996 (25.04.96)

(21) International Application Number: PCT/GB95/02457

(22) International Filing Date: 18 October 1995 (18.10.95)

(30) Priority Data:

9420989.7	18 October 1994 (18.10.94)	GB
9511729.7	9 June 1995 (09.06.95)	GB

(71) Applicant *(for all designated States except US)*: SCOTTISH CROP RESEARCH INSTITUTE [GB/GB]; Invergowrie, Dundee DD2 5DA (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: CHAPMAN, Sean, Nicholas [GB/GB]; 190 Lochee Road, Dundee DD2 2NF (GB). SANTA CRUZ, Simon, Peter [GB/GB]; Tayview, 29 Hazel Avenue, Dundee DD2 1QD (GB). OPARKA, Karl, John [GB/GB]; 2 Pipers Way, Auchterhouse, Dundee DD3 0QS (GB). WILSON, Thomas, Michael, Aubrey [GB/GB]; The Coach House, 4 Balruddery Meadows, Invergowrie DD2 5LJ (GB).

(74) Agent: MURGITROYD & COMPANY; 373 Scotland Street, Glasgow G5 8QA (GB).

(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHOD OF PRODUCING A CHIMERIC PROTEIN

(57) Abstract

A method of producing a chimeric protein from i.e. a plant virus coding for such a protein. The method allows the production of large (i.e. 25 kDa) proteins which assemble with the virus in infected host cells and are arranged on the outer surface of chimeric viruses. A vector for the production of biologically useful proteins in such a manner is also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

METHOD OF PRODUCING A CHIMERIC PROTEIN

This invention relates to a method of producing a chimeric protein, eg a biologically active protein such as an antibiotic peptide.

Typical antibiotic peptides include the marginins, 23 amino acid-long alpha-helical peptides, originally identified from frog skin, which have significant antibacterial activity; the defensins which combat bacteria, fungi and some enveloped viruses such as herpes simplex virus and HIV; and the protegrins which are 16-18 amino acid-long antibiotic peptides with strong biocidal activity.

The protegrins form part of an array of antibiotic peptides that are used by mammalian phagocytes to destroy invading pathogens through non-oxidative processes. Typically the protegrins include 4 cysteine residues and form a double-stranded β-sheet structure and show sequence similarity with the antibiotic defensin peptides that are also involved in phagocyte defence responses. The defensins are cationic, cysteine-rich peptides of 29 to 34 amino acids that are formed almost entirely of β-sheet structures and that

have been shown to have biocidal activity against bacteria, fungi and some enveloped viruses, including herpes simple virus and HIV. Both the protegrins and defensins are expressed in phagocytes as pre-pro-proteins which are cleaved to release the biocidal peptides from the carboxy-terminus of the protein.

Because of their antibacterial activity it may not be convenient to synthesize these antibiotic peptides by genetic engineering in conventional prokaryotic expression systems. Solution synthesis of large amounts of these peptides with a variety of amino acid modifications may be possible, but is not currently considered commercially viable, since a significant drop in yield occurs in the manufacture of peptides of over 25-30 amino acid residues.

Eukaryotic expression systems (yeast, insect, animal or plant cells which produce foreign proteins or peptides) may be necessary if there is a need for post-translational modification of the desired protein, but fermentation processes for such eukaryotic expression systems are expensive to maintain, provide little flexibility in terms of scaling the process up to industrial production levels and are very susceptible to contamination. Processing and purification of the desired protein can also be complex and costly.

The use of plants and benign plant viruses offers an opportunity to produce foreign proteins with minimal host cell contamination, thereby reducing contamination problems which could affect successful achievement of the required regulatory body approval for human or veterinary applications.

It has been proposed in WO92/18618 to use plant viruses

as vector systems for the expression of foreign nucleotide sequences. WO92/18618 describes the use of a Comovirus (Cowpea Mosaic Virus or CPMV) as an effective vector for such expression and also mentions other spheroidal viruses such as HIV and Picorna-viruses. *Picornaviridae* generally comprise particles of 22-30nm having cubic symmetry; *Comoviridae* have a pair of 28nm particles with a similar symmetry, and HIV is a member of the *Retroviridae* which are generally enveloped 100nm particles containing an icosahedral nucleocapsid.

One disadvantage of the system disclosed in WO92/18618 is that the geometry of the spheroidal viruses precludes large proteins from being produced, since the size and number of chimeric proteins per virus particle (generally 60 for icosahedral virus particles) is limited by the spheroidal geometry of the virus.

Construction of chimeric proteins in such viruses is also limited to the insertion of the foreign component into a loop in a native virus protein, eg the β-B to β-C loop in VP23 of CPMV, where such insertion does not affect the geometry of the coat protein and/or its ability to self-assemble into a virus particle (virion). As can be appreciated, the size of the peptide which can be tolerated in such an insertion is fairly limited; polypeptides of a maximum of 26 amino acids in length are cited by WO92/18618. Larger polypeptides present in internal insertion sites in coat or capsid proteins of the viruses exemplified may result in disruption of the geometry of the protein and/or its ability to successfully interact with other coat proteins leading to failure of the chimeric virus to assemble. Modified viruses which cannot self-assemble might not infect other host cells and produce

whole plant infection. This possible lack of ability to spread the infection of the modified virus constitutes a significant disadvantage in the prior system.

The present invention contemplates the use of benign high copy number rod-shaped viruses, preferably plant viruses such as potato virus X (PVX), to produce foreign protein connected to viral coat protein subunits. When assembled, the virus particles comprise long helical arrays of more than 1000 identical chimeric proteins (which are typically coat protein - foreign protein fusion molecules) per virion. Generally the foreign protein portion will be displayed on the outer surface of the virus particles.

A suitable proteolytic degradation site (eg elastase or CNBr) may be engineered into the chimeric protein to permit release of the foreign protein portion from purified virus material. Given the size of the foreign protein and the relevant composition of the possible viruses, it is estimated that between 10 and 30% of the total weight yield of virus particle could comprise the foreign protein. Release of the foreign protein by proteolytic cleavage can be a simple purification regime, followed by removal of the residual innocuous plant virus itself. Yields of plant virus up to 5g per kg wet weight of leaf from potato or tobacco are possible and hence the yields of foreign protein could be very substantial.

If the foreign protein is left attached to the chimeric protein in the virus particle, the whole virus particle can also be used as a vector for expression and presentation of peptide epitopes for vaccination of animals and/or the delivery of therapeutic single-

stranded RNA molecules. This may be of utility in the delivery of anti-sense or triplex nucleotides.

The present invention provides a method of producing a chimeric protein comprising:

a. providing a rod-shaped recombinant virus or pseudovirus containing a polynucleotide encoding a chimeric protein having a first (viral) portion and a second (non-viral) portion, the chimeric protein being capable of assembly into a virus particle such that the second portion is disposed on the exterior surface of the assembled virus particle;

b. infecting a host cell with the virus or pseudovirus; and

c. allowing replication of the virus or pseudovirus and expression of the chimeric protein in the host cell.

The term "rod-shaped" as applied herein to viruses includes filamentous or flexuous viruses, which are preferred. It is advantageous to use a virus which is flexuous (ie which can bend easily) since chimeric proteins with large second portions may be able to assemble more easily into virus particles (virions) which are flexuous than those which are rigid. PVX is preferred since it forms a flexuous virion.

The virus or pseudovirus can preferably assemble in the host cell to produce infective virus particles which comprise nucleic acid and chimeric protein. This enables the infection of adjacent cells by the infective virus or pseudovirus particle and expression

of the chimeric protein therein.

The host cell can be infected initially with virus or pseudovirus in particle form (ie in assembled rods comprising nucleic acid and protein) or alternatively in nucleic acid form (ie RNA such as viral RNA; cDNA or run-off transcripts prepared from cDNA) provided that the virus nucleic acid used for initial infection can replicate and cause production of whole virus particles having the chimeric protein.

The term "pseudovirus" as used herein means a virus-derived nucleic acid sequence optionally assembled into particles and having an incomplete viral genome as compared to wild-type virus but retaining sufficient viral genes to allow replication and assembly of the pseudovirus. The virus or pseudovirus may contain genetic material foreign to the wild-type virus.

Optionally, the virus or pseudovirus can be purified from the host cell in order to concentrate the chimeric protein, ie by polyethylene glycol precipitation and/or density gradient centrifugation.

Optionally, the method may include the step of separating a protein derived from the second portion from the remainder of the chimeric protein after the virus or pseudovirus has been purified from the host cell.

A linker peptide can be incorporated between the first and second portions and may have the function of spacing the two portions from one another, reducing stearic restrictions. Optionally the linker peptide may contain a proteolytic or chemical cleavage site.

The term "proteolytic or chemical cleavage site" refers to a short sequence of amino acids which is recognisable and subsequently cleavable by a proteolytic enzyme or chemical means. Suitable proteolytic enzymes include trypsin, pepsin, elastase and the like. Alternatively the proteolytic or chemical cleavage site may be a site which is vulnerable to cleavage by other means, for example by addition of chemicals such as cyanogen bromide (CNBr) or acids or by shear. Preferably, the proteolytic or chemical cleavage site is an elastase cleavage site, but other suitable proteolytic cleavage sites can be used with corresponding enzymes.

The protein derived from the second portion may be separated from the remainder of the chimeric protein before assembly of the virus particle, eg during expression of the genetic material coding for the chimeric protein, or during assembly of the chimeric protein into a virus particle. In this embodiment the host cell will contain free protein derived from the second portion. This embodiment can be useful when expression of very large proteins derived from the second portion is desired. In such an embodiment, the proteolytic or chemical cleavage site may be selected to cleave automatically in a virally-infected host cell.

The term "proteolytic or chemical cleavage site" may thus also include sequences that cleave automatically such as the FMDV (Foot and Mouth Disease Virus) 2A protease.

The proteolytic or chemical cleavage site may be an integral part of either the first or second portion. Hence either/or both of the portions may include an

integral proteolytic or chemical cleavage site.

Thus the present invention also provides a method of producing a chimeric protein as defined above, wherein the protein derived from the second portion is purified directly from the host cell after expression.

The second portion and/or the protein derived therefrom may be relatively large eg over 10kDa. Proteins of 25-30 kDa are suitable for production by the method and even proteins up to 60-70 kDa have been shown to be produced by the method of the invention.

The first (viral) portion of the chimeric protein may be any protein, polypeptide or parts thereof, derived from a viral source including any genetically modified versions thereof (such as deletions, insertions, amino acid replacements and the like). In certain embodiments the first portion will be derived from a viral coat protein (or a genetically modified version thereof). Mention may be made of the coat protein of Potato Virus X as being suitable for this purpose. Preferably the first portion has the ability to aggregate into particles by first-portion/first portion association. Thus, a chimeric protein molecule can assemble with other chimeric protein molecules or with wild-type coat protein into a chimeric virion.

In a preferred embodiment of the invention the particle is derived from a potyvirus or even more preferably a potexvirus such as PVX, and in such an embodiment, the second portion is preferably disposed at or adjacent the N-terminus of the coat protein. In PVX, the N-terminus of the coat protein is believed to form a domain on the outside of the virion.

The second portion of the chimeric protein may be any protein, polypeptide or parts thereof, including any genetically modified versions thereof (such as deletions, insertions, amino acid replacements and the like) derived from a source other than the virus from which the first portion is derived. In certain embodiments the second portion or the protein derived therefrom is a biologically active or useful molecule. The second portion or the protein derived therefrom may also be a diagnostic reagent, an antibiotic or a therapeutic or pharmaceutically active agent. Alternatively the second portion or the protein derived therefrom may be a food supplement.

In an alternative embodiment, the second portion or the protein derived therefrom may be an indicator protein chosen for its ability to indicate the location of the chimeric protein or of the virus particle. Such an example is the 25kDa jellyfish green fluorescent protein.

The polynucleotide coding for the second (non-viral) portion may be inserted into an appropriate restriction site in the viral genome. The restriction site adopted for such insertion may be naturally occurring in the viral genome or artificially constructed therein and the polynucleotide coding for the second portion may be ligated therein by conventional means. General techniques for cloning of foreign nucleic acid and construction of chosen restriction sites is comprehensively described in the art and is within the scope of the skilled person.

It is preferred that the polynucelotide coding for the second portion is inserted at or adjacent a terminus of the polynucleotide coding for the first portion, such

that upon translation the chimeric protein has the first portion at one end and the second portion at the opposite end. It is not necessary for the first portion to comprise a whole virus coat protein, but this remains an option.

The virus particle may be formed by the assembly of chimeric proteins only or by the mixed assembly of chimeric proteins together with some unmodified or less modified forms of the naturally occurring wild-type coat protein which forms the basis of the first portion. For a mixed virus particle of the latter type, there must be present polynucleotide(s) encoding the chimeric protein and the naturally occurring coat protein. The appropriate protein-coding sequences may be arranged in tandem on the same molecule. An alternative would be co-infection (for example of mutually dependant defective viruses or pseudoviruses) of two or more viruses or pseudoviruses, or infection by chimeric virus of a host cell or whole organism (such as a plant) which expresses such a protein intrinsically.

An advantage is gained by using a virus which forms a particle with a relatively high pitch of helix. PVX has a pitch of 3.4nm and is to be preferred over viruses with a lower pitch. Virus particles with higher pitches may be able to accommodate larger protein insertions on their surfaces since their coat proteins assemble with more space between them than coat proteins of viruses with lower pitches.

A virus or pseudovirus genetically modified to express the chimeric protein forms a further aspect of the present invention, as does any host cell infected with such a virus or pseudovirus.

Preferably, the host cell used to replicate the virus or pseudovirus is a plant cell where the virus is a plant virus, although insect cells, mammalian cells and bacteria can be used with viruses which will replicate in such cells.

While modifications and improvements may be incorporated without departing from the scope of the invention, embodiments will now be described by way of the following examples and with reference to the accompanying drawings in which:

Fig 1a shows the structure of a gene for a chimeric protein and of the overcoat vector pTXS.L2a-CP for use in the present invention;
Fig 1b is a schematic diagram showing the major features of plasmids useful in the methods of the present invention;
Fig 2 shows a western blot of wild type and chimeric protein taken from leaves of a plant infected by a wild-type and a chimeric virus;
Fig 3 a, b, c and d show leaves of plants infected with recombinant virus;
Fig 4 a, b, c, d and e are micrographs illustrating the subcellular distribution of chimeric protein expressed from chimeric virus nucleic acid;
Fig 5 is an electron micrograph showing aggregation and immuno-gold labelling of chimeric viruses;
Fig 6 a, b and c are electron micrographs of negatively-stained chimeric viruses; and
Fig 7 is a photograph of a *N benthamiana* leaf systemically infected with a chimeric virus.

EXAMPLE 1

A general strategy for the production of large quantities of recombinant proteins is given below using PVX as an example. A similar strategy could be employed for other flexuous filamentous or rod-shaped viruses. A cDNA clone of potato virus X is first modified to produce fusion proteins between the viral coat protein and proteins with biological activity or other commercial applications. The feasibility of this approach has been demonstrated as described below by creating a translational fusion between the green fluorescent protein (25 kDa) of *Aequorea victoria* (1) and the PVX coat protein (also around 25 kDa). Functional chimeric viruses have also been made which are able to express recombinant genes encoding fusions between the PVX coat protein and the kanamycin resistance protein Neomycin phosphotransferase (25 kDa) and between PVX coat protein and the more complex enzymes β-galactosidase (10-13 kDa) and β-glucuronidase (68 kDa) respectively.

The green fluorescent protein (GFP) from *A. victoria* (1) is a reporter of gene expression in heterologous systems (3-6). GFP has an advantage over other marker proteins in that it can be detected non-invasively, without any requirement for exogenous substrates or co-factors (3) since it fluoresces intrinsically without a requirement for exogenous substrate. In addition, fluorescence of GFP is retained in fusion proteins allowing the subcellular localization of fusion proteins (4).

PCR-mutagenesis of a full-length cDNA copy of the potato virus X genome can be performed to create a synthetic coding sequence comprising the gene coding for the protein of interest, the foot and mouth disease virus 2A protease gene, and the potato virus X coat

protein gene. The PVX genome is contained within the known plasmid pTXS (Fig. 1, reference 25).

When reassembled the modified cDNA copy of the viral genome can be used as a template to synthesize *in vitro* run-off transcripts. Inoculation of transcripts to plants can be performed by manual abrasion of carborundum coated leaves of either *Nicotiana clevelandii* or *N benthamiana*.

When the above approach was followed using PVX modified to express GFP-CP fusion protein, between two and three days post inoculation the presence of fluorescent regions in the virus infected plants could be observed by eye on inoculated leaves by viewing plants under ultraviolet light. At about ten days post inoculation GFP-mediated fluorescence was detected in systemic (non-inoculated) leaf tissue (Figure 7). This fluorescence was specific to the green fluorescent protein and was not observed on control plants inoculated with wild-type PVX.

Electron microscopic analysis of viral particles showed a clear increase in particle width in plants infected with the GFP-CP containing virus compared with particles isolated from plants infected with wild-type PVX (Figure 6).

In the strategy used above, foreign proteins were expressed by fusing them to the amino-terminus of the PVX coat protein. However other sites may be possible, eg carboxy-terminus surface loops on some other rod-shaped or filamentous viruses.

Data from previous studies suggest that fusion of the proteins to the amino terminus of the PVX coat protein

is most likely to be successful. Biochemical, immunological and tritium bombardment data suggest a model for the structure of the PVX coat protein (10) in which the N-terminal 33 amino acids form a domain of β-sheet on the outside of the virion. In contrast, the C-terminus of the PVX coat protein, which also forms part of a β-sheet structure, is inaccessible from the outside of the virion and deletions within it do not permit the virus to infect plants systemically.

As an additional optional strategy, the foot and mouth disease virus (FMDV) 2A protease sequence (12) can be positioned between the foreign and coat protein sequences. The FMDV 2A protease is a short (19 amino acid) peptide which acts *in cis* to cleave the FMDV polyprotein in a co-translational mechanism. This protease has been shown to effect the cleavage of synthetic polyproteins both *in vitro* and *in vivo* (13). The inclusion of the 2A protease sequence between the GFP and coat protein can generate a mixed pool of fusion and cleaved proteins in virus infected cells. The presence of free coat protein, generated by 2A protease mediated cleavage, may circumvent this problem by allowing assembly of virions composed of both free (ie cleaved) and fused coat protein subunits.

The formation of virions is an absolute requirement of PVX for systemic infection of plants (15). The demonstration herein that GFP-coat protein fusions do assemble into virions (Fig 7) and spread indicates that the size of GFP (25kDa) does not interfere with virion assembly. Fusion proteins which fail to assemble due to size or other constraints can be produced in constructs carrying the FMDV 2A protease, or in plants which are modified to express wild-type coat protein for the particular virus used. The sequence of the 2A

protease peptide can be modified to increase or decrease the efficiency of co-translational cleavage.

EXAMPLE 2

This example describes a modified form of PVX which expresses a chimeric gene encoding a fusion between the *Aequorea victoria* green fluorescent protein and the PVX coat protein and assembles into virions that are over twice the diameter of wild-type PVX. The modified virus moves from cell-to-cell and systemically. The example demonstrates the potential of fusions between non-viral protein and virus coat protein for production of high levels of non-viral proteins in plants.

The plasmids used in this work were derived essentially from the plasmid pTXS which contains the PVX genome and a T7 promoter (described in 25). Fig 1b shows the following main features of the plasmids: the virus RNA-dependent RNA polymerase gene (RdRp); virus genes encoding movement proteins (M1, M2, M3); the virus coat protein gene (CP); promoters from T7 bacteriophage (T7) or for the 35S RNA of CaMV (CaMV35S); the transcriptional terminator of the nopaline synthase gene of *Agrobacterium tumofaciens* and various restriction enzyme sites.

The plasmid pCXA3 was constructed by transfer of the PVX cDNA from pTXS into the plasmid pB1220.5 between the CaMV 35SRNA promoter and the nopaline synthase gene terminator. The plasmid pB1220.5 is similar to the plasmid pB1221.1 but without the GUS gene (described in 27). The junction between the promoter and the PVX cDNA was modified by oligonucleotide directed mutagenesis to the sequence (5′)gatttggagagga*gaaaactaaacca(3′) in which * denotes the most 3′ non-transcribed position in the promoter

sequence and the most 5′ transcribed position in the viral genome (28). Construction of the pVX201 vector from pCXA3 and pPC2S exploited unique restriction sites at positions 4945 (Apa1) and 6302 (Xho1) of the PVX cDNA (25).

GFP cDNA was PCR-amplified with primers (5′)gccaatcgatcatgagtaaaggag(3′) on the positive strand and (5′)ggaagtcgacacatttatttg(3′) from the negative strand. The bold type represents the initiation and termination codons of the GFP gene (29). The underlined type represents Cla1 and Sal1 sites used to introduce the PCT-amplified sequence into pPVX201 to generate pPVX204. The plasmid pTXS.GFP was made by substitution of the region of pPVX204 containing the GFP sequence into the homologous region of pPC2S.

The plasmid pTXS.GFP carries a full-length cDNA copy of the potato virus X (PVX) genome into which the GFP gene has been inserted. Inoculation of plants with transcripts synthesized *in vitro* from pTXS.GFP results in the expression of free GFP in infected cells (5). We prepared a derivative of pTXS.GFP, pTXS.GFP-CP, to create a translational fusion between the carboxyterminus of the GFP and the amino-terminus of the PVX coat protein (CP). pTXS.GFP was used as a template to produce the GFP-2A-CP fusion gene by overlap extension PCR using flanking oligonucleotides complementary to the PVX genome and mutagenic oligonucleotides to incorporate the 2A protease coding sequence. Amplified product was subcloned into pTXS.GFP as a 1.5 kbp fragment using the unique restriction sites Cla1 and Xho1 to give pTXS.GFP-CP. Fig. 1a shows a schematic representation of viral cDNAs used to synthesize infectious run-off transcripts for the GFP-2A-CP fusion gene. The predicted Mrs of the

four viral proteins common to all constructs are indicated (K=kD). The polypeptide chain lengths of the CP, GFP and 2A protease (2A) enclosed by the constructs are shown. The bars indicate the position of the subgenomic promoter for the CP. TXS=wild-type PVX; TXS.GFP=PVX modified to express free GFP from a duplicated subgenomic promoter; TXS.GFP-CP=PVX modified to express the GFP-2A-CP fusion protein.

Because the GFP and PVX CP are of similar sizes, having molecular weights of 26.9 kD and 25.1 kD respectively, it was expected that in a homogenous population of fusion protein steric effects would prevent virion formation. Assembly of fusion protein into virions might be facilitated by the presence of a pool of free CP. Therefore the GFP and CP nucleotide sequences in pTXS.GFP-CP were separated by sequence coding for sixteen amino acids from the foot-and-mouth disease virus (FMDV) 2A peptide. The 2A region of FMDV mediates a primary (co-translational) processing event between the 2A and 2B regions of the FMDV polyprotein (12) that results in inhibition of peptide bond formation (13).

In vitro run-off transcripts (14), synthesized from pTXS.GFP and pTXS.GFP-CP (plasmids were linearized with Spe 1 prior to *in vitro* transcription reactions as described in reference 14), were infectious when inoculated to plants; virus derived from transcript-infected plants is subsequently referred to as PVX.GFP and PVX.GFP-CP respectively.

Following inoculation of either *Nicotiana clevelandii* or *N. benthamiana*, both PVX.GFP and PVX.GFP-CP caused the development of green fluorescent regions which were first detectable by eye under UV illumination between

two and three days post inoculation (Fig. 3A, C). Subsequent long-distance movement of the virus to developing leaves led to the appearance of green fluorescence in systemically infected leaves (Fig. 3B, D). The rate at which fluorescent regions spread on inoculated leaves was slower in PVX.GFP-CP infected plants than PVX.GFP infected plants and the appearance of fluorescence in systemically infected leaves was delayed in plants infected with PVX.GFP-CP compared with PVX.GFP infected plants.

Fig. 3 shows leaves of *N. benthamiana* infected with either PVX.GFP (A, B) or PVX.GFP-CP (C,D). Leaves were viewed under UV illumination (365 nm) generated from a Blak Ray B100-AP lamp (Ultra-Violet Products) and photographed using a Wratten 58 filter to eliminate chlorophyll auto-fluorescence. The pattern of virus spread in both cases is similar. A and C identify inoculated leaves showing the development of characteristic circular legions. B and D identify systemically infected leaves showing fluorescence associated predominantly with the leaf veins. The developing leaf (D) was undergoing the sink-source transition (20) resulting in lack of virus movement into the apical portion of the leaf.

Fig 4a is a confocal fluorescence image of a systemically infected leaf in transverse section showing the location of PVX.GFP-CP containing viroplasms within individual cells of the leaf. 4b is a bright field image of section shown in (A) showing the typical arrangement of epidermis (E), palisade (P) and mesophyll(M) cells. A vascular bundle (B) is also present (scale=50 μm). 4c is a confocal image of palisade cells from a leaf systemically infected with PVX.GFP-CP showing the GFP-containing viroplasms (V)

assembled into cage-like structures (scale=5 μm). 4d shows a leaf trichome systemically infected with PVX.GFP, in which the GFP is associated with the nucleus (N) and the cytoplasm. 4e shows a leaf trichome systemically infected with PVX.GFP-CP, in which the GFP is predominantly targeted to viroplasms (V) within individual trichome cells (scale=10 μm).

In systemically infected (ie non-inoculated) leaves both PVX.GFP and PVX.GFP-CP moved from the phloem into surrounding bundle sheath and mesophyll cells and eventually into the epidermis (Fig. 4A, B). Under the confocal microscope transverse sections of the systemically infected leaves showed that in PVX.GFP-CP infected cells green fluorescence was detected predominantly in viroplasms, cytoplasmic structures comprising aggregated viral particles that often appeared as continuous cage-like structures within the cell (Fig 4C, 5). By contrast, in PVX.GFP infected cells, the green fluorescence was associated with nuclei and showed a relatively uniform distribution throughout the cytoplasm. This difference in the subcellular distribution of the GFP was seen clearly in leaf trichome cells (Fig. 4D, E).

The distribution of fluorescence suggested that the majority of GFP produced in PVC.GFP-CP infected plants was still fused to the CP and that these fusion proteins were assembling into virions, which subsequently formed viroplasms.

Western blotting of protein extracts from inoculated *N. clevelandii* leaves, probed with CP specific antiserum (16), showed that most of the immunoreactive protein in PVX.GFP-CP infected plants comprised the fusion protein. Protein extracts were prepared by grinding

leaf tissue in two volumes (w/v) protein extraction buffer (15). An equal volume of 2x SDS load buffer was added and the extracts were boiled for two minutes. Proteins were electrophoresed, blotted to nitrocellulose and probed with rabbit polyclonal anti-PVX CP antiserum as described previously (16) .Fig 2 illustrates the data obtained. Protein was prepared from mock inoculated control plants (lane 2), or from plants inoculated with in vitro transcripts synthesized from plasmid DNAs (TXS=lane 1; TXS.GFP-CP=lane 3; TXS.GFP=lane 4). Mrs of native CP, the GFP-2A-CP fusion protein and CP released by 2A protease mediated cleavage are 25.1, 53.2 and 24.8 kD respectively. The Mrs of standards are shown to the left of Fig 2 in kD.

The low level of smaller immunoreactive protein detected in PVX.GFP-CP infected tissue is assumed to result from processing of the fusion protein mediated by the FMDV 2A peptide rather than from contamination with virus deletion mutants as similar ratios of fusion to free protein were observed in all other samples analyzed and RT-PCR analysis of the same samples used for protein analysis showed no evidence of deleted forms of the viral genome (17). In addition when blots were probed with GFP specific antiserum the ratio of free protein to fusion protein was the same as that observed using anti-CP antiserum (17).

In order to determine the subcellular location of the viral CP ultrathin sections of inoculated leaves were prepared for immuno-gold labelling, using a polyclonal antibody to the PVX CP. Leaf tissues were fixed and embedded in Araldite (TM) resin for immuno-gold labelling as described previously (17). Ultrathin sections on nickel grids were labelled using polyclonal rabbit antiserum to the PVX CP followed by goat anti-

rabbit gold conjugate (GAR-15 nm, Amersham International). Aggregation of the filamentous virions into viroplasms is marked with arrows in Fig 5. Dense gold labelling was predominantly associated with the viroplasms in both PVX.GFP and PVX.GFP-CP infected cells. The pattern of virus aggregation seen in the electron microscope for both PVX.GFP-CP (Fig. 5) and PVX.GFP was remarkably similar to the cages of viroplasm seen with PVX.GFP-CP under the confocal microscope (Fig. 4c).

For negative staining, virus particles were trapped from virus infected sap extracts by immuno-sorbent electron microscopy (18) using anti-PVX CP antiserum, and stained with 2% sodium phosphotungstate (pH 7). Analysis of negatively stained virus samples under the electron microscope revealed that PVX.GFP-CP virions were decorated along their length with globular extensions (Fig. 6a,b). Fig 6c shows negatively stained virus rods isolated from PVX.GFP infected plants (scale=50 nm). Differences in virion diameter are seen most clearly where virions are aligned in parallel (a and c, large darts). In Fig 6b small globular extensions (small darts) are apparent along the length of the PVX.GFP-CP virus (scale=25 nm). The PVX.GFP-CP virions had a mean diameter of 29.7 nm, more than twice the diameter of PVX.GFP virions (12.6 nm; Fig. 6c).

A modified form of PVX.GFP-CP, in which the FMDV 2A peptide sequence carries three amino acid substitutions, introduced to prevent processing of the polyprotein, was unable to move from cell-to-cell and did not give rise to fluorescent viroplasms. Infections with this mutant were restricted to single epidermal cells and fluorescence was detected uniformly

throughout the cytoplasm and in association with nuclei, as observed for PVX.GFP infections (17), suggesting that the presence of free CP is essential for either initiation of elongation of virions.

The fluorescence generated by the GFP attached to virions was intense, allowing rapid detection of viral aggregates within individual living cells. Furthermore, confocal microscopy allowed the noninvasive imaging of the pathway of cell-to-cell movement of virus-GFP constructs, pinpointing the specific cell types in which virus accumulated. For confocal imaging leaves were excised from the plant and sectioned transversely into 200 μm slices using a vibrotome. The sections were immediately mounted in water and viewed under a Bio-Rad MRC 1000 confocal laser scanning microscope at an excitation wavelength of 488 nm using a krypton-argon laser.

Previous descriptions of assembly competent plant RNA viruses carrying CP extensions have involved small oligopeptide fusions (19). The data presented in this example suggest that the system described could be used for the production of proteins that are at least as large as the viral CP of PVX.

The strategy described to generate GFP-coat protein fusions can be easily applied to proteins other than GFP. We modified the plasmid pTXS.GFP-CP which carries the GFP-2A-CP fusion protein gene to enable the facile insertion of novel coding sequence as a fusion to the 2A-CP cassette. This modified plasmid, pTXS.L2a-CP shown in Fig. 1a (deposited under No NCTC 12918 at the National Collection of Type Cultures at 61 Colindale Avenue, London NW9 5HT on 18 October 1995) carries a series of unique restriction enzyme recognition sites

(Cla1, Ega1, Sma1, Ehe1) or polylinker that replaces the GFP coding sequence of pTXS.GFP-CP. By digesting the vector pTXS.L2a-CP at one or more of the polylinker restriction enzyme sites it is possible to insert the coding sequence for any given protein such that a fusion protein gene is created comprising the novel gene, the FMDV 2A peptide and the PVX coat protein as a translational fusion.

The plasmid vector pTXS.L2a-CP was prepared by PCT-based mutagenesis of the plasmid pTXS.GFP-CP using standard techniques (26). The oligonucleotide 2aL5' was annealed to the primer 2aL3' and extended with T4 DNA polymerase.

The sequence of primers used was

```
2aL5': 5'  TCG GCC GTC CCG GGG GCG 3'
           ||| ||| ||| ||| ||| |||
2aL3': 3'  AGC CGG CAG GGC CCC CGC GGT TAA AAC TGG AAG

           AAT TCG AAA 5'
```

The extended product was gel purified and cloned into the plasmid M13RK8.2 (30). An Eag 1/ Afl 11 fragment was excised from the resulting plasmid and cloned between the same sites of the plasmid pTXS.GFP-CP in place of the GFP gene.

Thus, the nucleotide sequence of the new linker in pTXS.L2a-CP is

```
       Cla1      Eag1       Sma1     Ehe1

      /_____\  /_____\   /_____\ /____\
Nts:  AT CGA TCC GGC CGT CCC GGG GGC GCC AAT TTT
Amino acids:              Pro Gly Gly Ala Asn Phe
```

Insertion of foreign genes into the pTXS.P-CP polylinker are most easily performed by PCR amplification of the foreign gene using oligonucleotides designed to incorporate appropriate restriction enzyme recognition sites at the 5'- and 3'-termini of the foreign coding sequence such that the gene for the synthetic polyprotein comprises a single open reading frame. We have demonstrated the utility of this approach using the gene encoding neomycin phosphotransferase (NPT) which confers resistance to the antibiotic kanamycin and is present in most commercially available plasmids as a selection tool. The 0.73 kb (NPT) coding sequence was inserted into the polylinker of pTXS.P-CP to give the plasmid pTXS.NPT-CP. Transcripts synthesized in vitro from the pTXS.NPT-CP template were infectious on plants and the virus moved both locally and systemically. Assembly of PVX.NPT-CP virions results in "overcoat" virus particles carrying the NPT protein on the surface of the virions.

The advantages of the invention are as follows:

(i) Standard purification procedures exist (eg polyethylene glycol precipitation and centrifugation) for these highly stable virus particles to remove plant proteins and cellular debris and to give an extremely pure suspension of plant virus particles. Plant viruses are innocuous to humans, ingestion experiments have already revealed that they pass straight through the intestine undamaged.

(ii) By attaching the foreign protein to each (or a subset of) coat protein subunits optionally with a suitable cleavage-sensitive linker sequence will allow, following virus purification from the infected plant

sap, foreign protein to be released into free solution simply by incubation with the appropriate proteolytic enzyme. The released virus particles remain stable and of high molecular weight so that they can be separated from the short peptide either by simple dialysis procedures (continuous flow type), or by differential centrifugation or selective precipitation.

(iii) Yields of cleaved foreign protein from such a system could reach 50% or more of the total weight of virus recovered. Each helical virus particle has 95% of its weight as coat protein, and each coat protein subunit has a molecular weight of approximately 25 kD. In the model system already developed the green fluorescent protein also has a molecular mass of approximately 25 kD. Yields of potato virus X can be extremely high (up to 5 gm/kg wet weight of infected leaf after several weeks).

(iv) The flexibility of scale that can be achieved in plants is also attractive in terms of reducing the cost of protein production and avoids the need for high level capital investment such as in animal or microbial cell culture facilities.

(v) The use of set-aside land and/or discredited crops such as tobacco for the alternative production of highly prized, pharmaceutically active proteins would lead to considerable added value in the peri-agricultural sector.

REFERENCES

The following documents referred to in the text are incorporated herein by reference:

1. B.L. Epel, Plant Mol. Biol. 26, 1343-1356 (1994): W.J. Lucan & R.L. Filbertson, Annu. Rev. Phytopathol. 32, 387-411 (1994); B.G. McLean, W. Waigmann, V. Citovsky & P. Zambryski, trends Microbiol. 1, 105-109 (1993).

2. D. C. Prasher, V.K. Eckenrode, W.W. Ward, F.G. Prendergast & M.J. Cormier, Gene 111, 229-233 (1992).

3. M Chalfie, Y. Tu, G. Euskirchen, W.W. Ward & D.C. Prasher, Science 263, 802-805 (1994).

4. S. Wang & T. Hazelrigg, Nature 369, 400-403 (1994).

5. D.C. Baulcombe, S. Chapman & S. Santa Cruz, Plant J.7, 1045-1053 (1995).

6. R.P. Niedz, M.R. Sussman & J.S. Satterlee, Plant Cell Rep. 14, 403-406 (1995).

7. S. Chapman, T. Kavanagh & D.C. Baulcombe, Plant J. 2, 549-557 (1992).

8. H.B. Scholthof, T.J. Morris & A.O. Jackson, Mol. Plant-Microbe Interact. 6, 309-322 (1993).

10. L.A Baratova et al., Virology 188, 175-180 (1992).

11. R. Koenig & L. Torrance, J. Gen. Virol 67, 2145-2151 (1986).

12. M.D. Ryan, A.M.Q. King & G.P. Thomas, J. Gen. Virol. 72, 2727-2732 (1991).

13. M.D. Ryan and J. Drew, EMBO J. 13, 928-933 (1994).

14. T. Kavanagh et al., Virology 189, 609-617 (1992).

15. S. Chapman, G. Hills, J. Watts & D.C. Baulcombe, Virology 191, 223-230 (1992).

16. C. Davies, G. Hills & D.C. Baulcombe, virology 197, 166-175 (1993).

17. C. Fasseas, I.M. Roberts & A.F. Murant, J. Gen. Virol. 70, 2741-2749 (1989).

18. I.M. Roberts, in Electron microscopy of proteins: 5. Viral structure, J.R. Harris & R.W. Horne, Eds. (Academic Press, New York, 1986), pp 293-357.

19. H Hamamoto et al., BIO/TECHNOLOGY 11, 930-932 (1993); R. Usha et al., Virology 197, 366-374 (1993); C. Porta et al., Virology 202, 949-955 (1994); T.H. Turpen et al., BIO/TECHNOLOGY 13, 53-57 (1995).

20. R. Turgeon, Annu. Rev. Plant Physiol. Plant Mol. Biol. 40, 119-138 (1989).

21. Abouhaidar, M.G. and Lai, R. (1989) J. Gen. Virol. 70, 1871-1875.

22. Kokyakov, V.N. Harwig, S.S.L., Panyutich, E.A., Shevchenko, G.M.A., Shamova, O.V., Korneva, H.A. and Lehrer, R.I. (1993) FEBS letters 327, #2, 231-236.

23. Pardi, A., Zhang, X-L., Selsted, M.E., Skalicky,

J.J. and Yip, P.F. (1992) Biochemistry 31, 11357-11364.

24. Lehrer, R.J., Ganz, T. and Selsted, M.E. Defensins: Endogenous Antibiotic Peptides of Animal Cells. (1991) Cell 64 229-230.

25. Kavanagh, T Goulden, M Santa Cruz, S., Chapman, S., Barker, I and Baulcombe, D.C. (1992) Virology, 189, 609-617.

26. Sambrook et al (1989. Molecular cloning, a laboratory manual. Cold Spring Harbour Laboratory.

27. Jefferson et al (1987) Plant Vol. Biol. Pep. 5, 387-405.

28. Guilley et al (1982) Cell 30, 763-773.

29. Prasher et al (1992) Gene 111, 229-233.

30. Waye et al (1985) Nucleic Acids Res. 13(23) 8561-8571.

Claims:

1 A method of producing a chimeric protein, the method comprising:

a providing a rod-shaped recombinant virus or pseudovirus containing a polynucleotide encoding a chimeric protein having a first (viral) portion and a second (non-viral) portion, the chimeric protein being capable of assembly into a virus particle such that the second portion is disposed on the exterior surface of the assembled virus particle;

b infecting a host cell with the virus or pseudovirus; and

c allowing replication of the virus or pseudovirus and expression of the chimeric protein in the host cell.

2 A method according to claim 1, wherein the chimeric protein assembles into a virus particle.

3 A method according to claim 1 or claim 2, wherein the virus or pseudovirus is subsequently purified from the host cell.

4 A method according to claim 2 or claim 3, including the step of cleaving the second portion or a protein derived therefrom from the first portion after purification of the virus or pseudovirus from the host cell.

5 A method according to any preceding claim, wherein

a linker peptide is incorporated between the first and second portions.

6 A method according to any preceding claim, wherein a proteolytic cleavage site is incorporated on one of or between the first and second portions.

7 A method according to claim 1, wherein the first and second portions are separated from one another before or during assembly of the virus particle, such that the host cell contains free protein derived from the second portion.

8 A method according to any preceding claim, wherein protein derived from the second portion is purified from the host cell after replication.

9 A method according to any preceding claim, wherein the virus or pseudovirus is derived from a plant virus.

10 A method according to any preceding claim, wherein the virus or pseudovirus is derived from potato virus X.

11 A method according to any preceding claim, wherein the second portion is disposed at or adjacent the N-terminus of the viral coat protein.

12 A method according to any preceding claim, wherein the second portion is a diagnostic reagent, an antibiotic, a therapeutic or pharmaceutically active agent, a vaccine or a food supplement.

13 A method according to any preceding claim, wherein the virus or pseudovirus particle comprises a mixture of chimeric protein and wild-type coat protein.

14 A method according to any preceding claim, wherein the virus or pseudovirus particle has a relatively high pitch of helix.

15 A method according to claim 12, wherein the pitch of the helix is more than 2nm.

16 A method according to any preceding claim, wherein the virus or pseudovirus is flexuous.

17 A method according to any preceding claim, wherein the host cell is infected with virus or pseudovirus in particle form.

18 A method according to any one of claims 1-16, wherein the host cell is infected with virus or pseudovirus in nucleic acid form.

19 A method according to any preceding claim, wherein the second portion or a peptide derived therefrom has a molecular weight in excess of 10 kDa.

20 A virus or pseudovirus genetically modified to express a chimeric protein, the chimeric protein having a first (viral) portion linked to a second (non-viral) portion, the chimeric protein being capable of self-assembly into a virus particle so that the second portion is disposed on the exterior surface of the assembled virus particle.

21 A host cell, plant, animal or insect infected with a virus or pseudovirus according to claim 20.

22 A polynucleotide capable of producing a virus or pseudovirus according to claim 20.

23 A chimeric protein produced by a method according to any one of claims 1-19.

24 The plasmid pTXS.L2a-CP as deposited under No NCTC 12918 on 18 October 1995 at the National Collection of Type Cultures.

FIGURE 1a

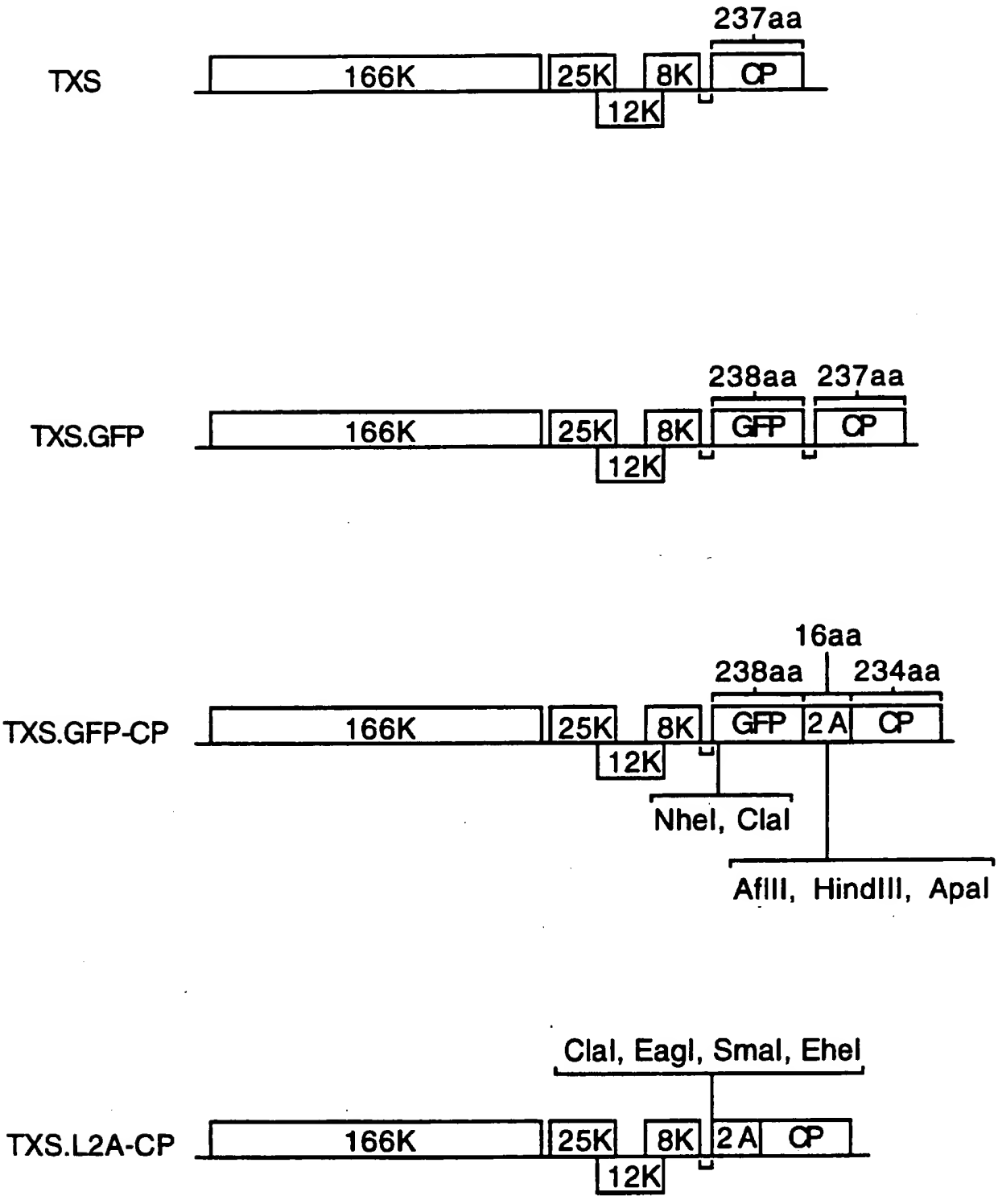
237aa
TXS
166K
25K
8K
CP
12K
238aa
237aa
TXS.GFP
166K
25K
8K
GFP
CP
12K
16aa
238aa
234aa
TXS.GFP-CP
166K
25K
8K
GFP
2A
CP
12K
NheI, ClaI
AflII, HindIII, ApaI
ClaI, EagI, SmaI, EheI
TXS.L2A-CP
166K
25K
8K
2A
CP
12K

FIGURE 1b

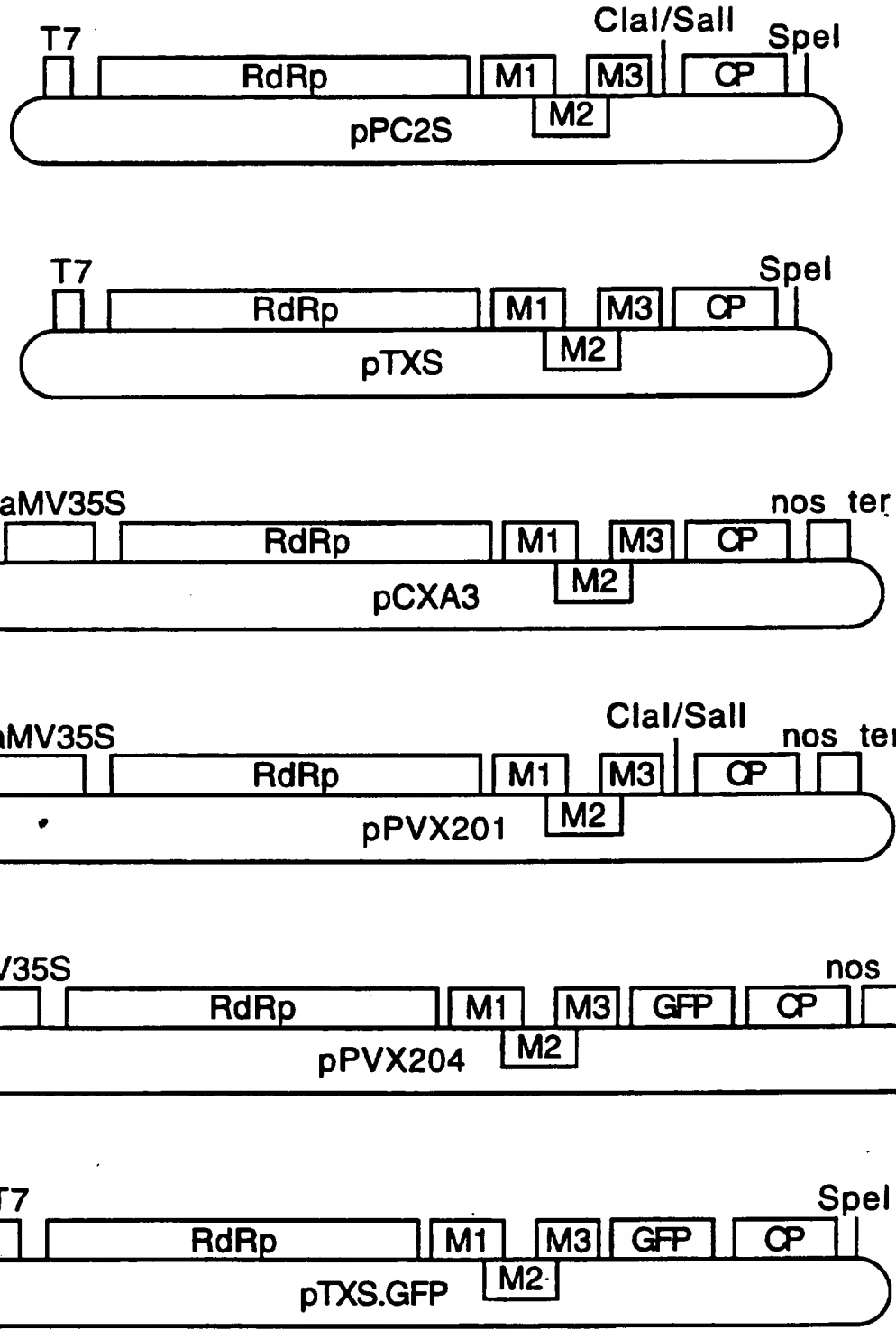
T7
RdRp
M1
M3
ClaI/SalI
CP
SpeI
M2
pPC2S
T7
RdRp
M1
M3
CP
SpeI
M2
pTXS
CaMV35S
RdRp
M1
M3
CP
nos ter
M2
pCXA3
CaMV35S
RdRp
M1
M3
ClaI/SalI
CP
nos ter
M2
pPVX201
CaMV35S
RdRp
M1
M3
GFP
CP
nos ter
M2
pPVX204
T7
RdRp
M1
M3
GFP
CP
SpeI
M2
pTXS.GFP

FIGURE 2

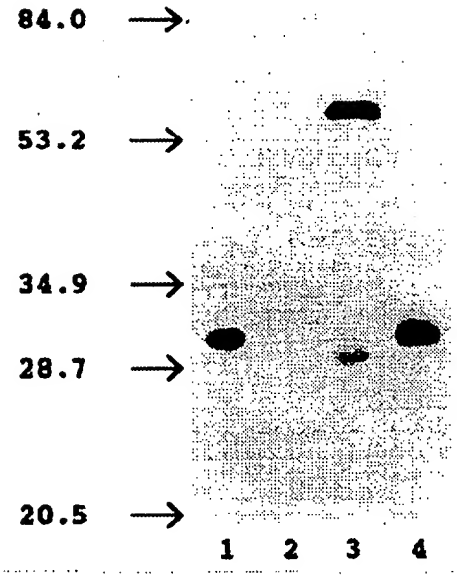

FIGURE 3

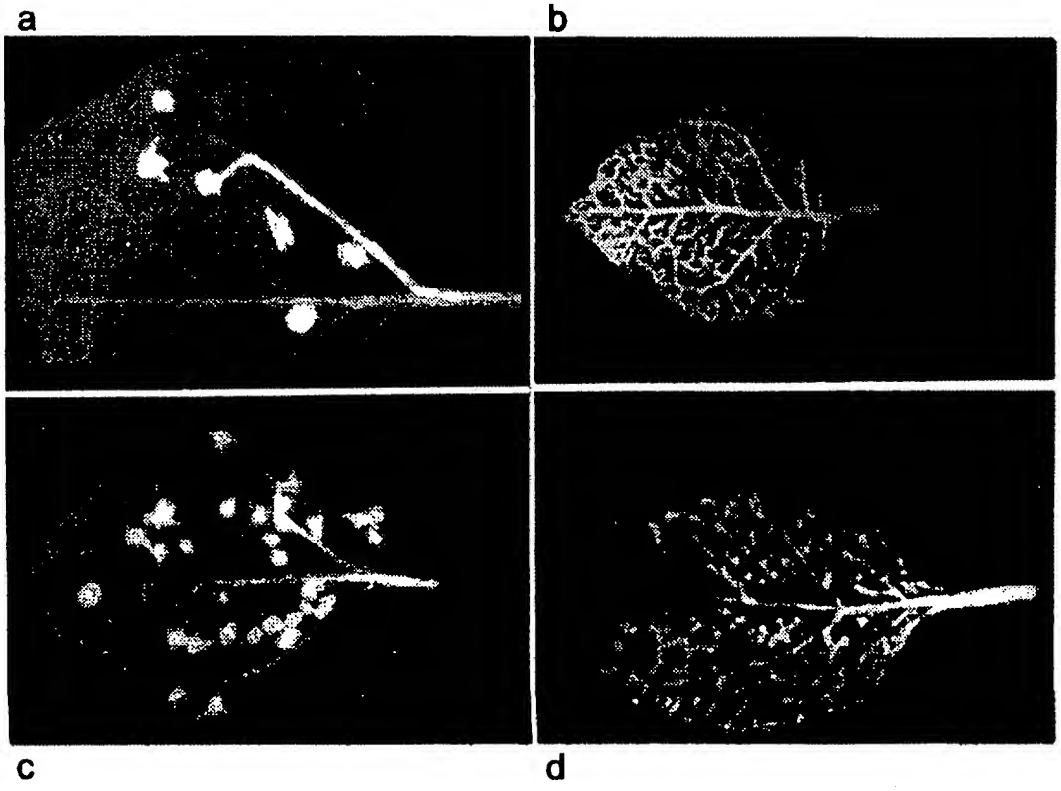

FIGURE 4

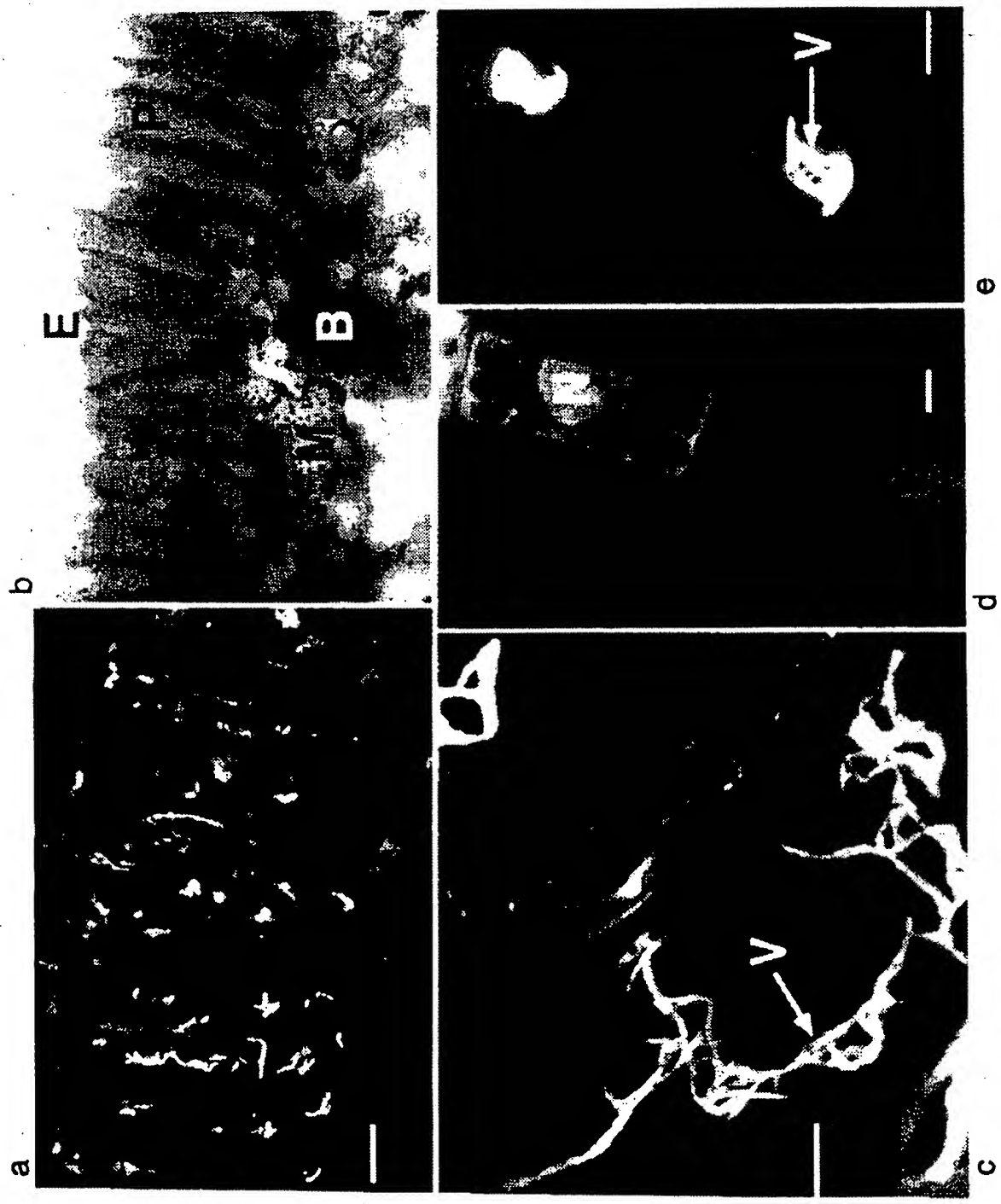

FIGURE 5

FIGURE 6

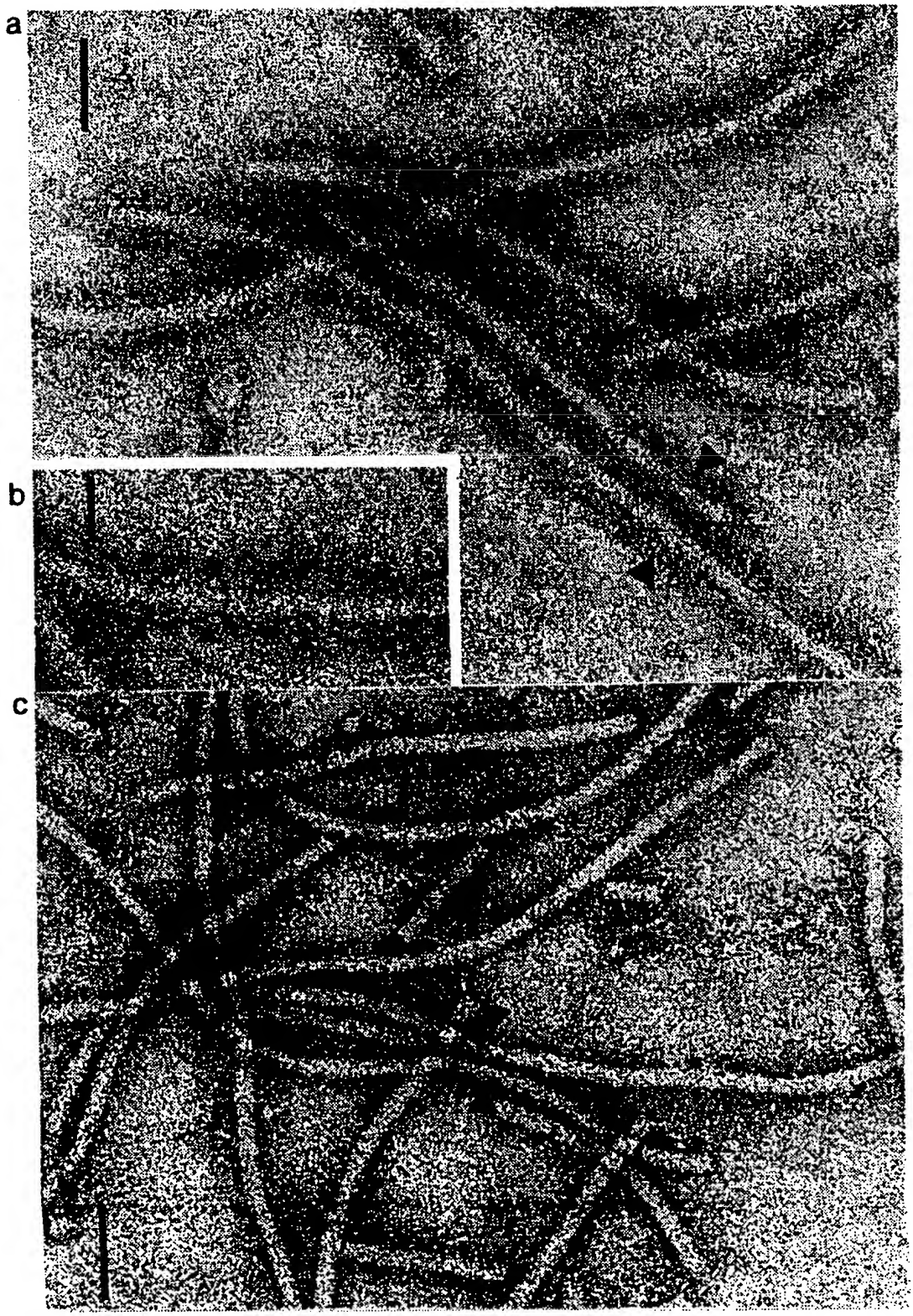

FIGURE 7